大数据技术为广播的长足发展提供技术支撑

摘　要：本文提出了大数据技术概念、特点及对广播媒体的重大意义，介绍了大数据技术管理平台的构建方式及采用的关键技术，探讨了大数据技术原理及大数据在安全方面遇到的问题。结合电台的工作实际，详细叙述了大数据技术在提高广播节目制作质量、精准的收听率调查、信息获取渠道及发挥主流媒体作用等方面的作用。

关键词：大数据；广播媒体；云计算；数据分析

中图分类号：TP391　　　　**文献标识码：**A

文章编号：1671–0134（2017）10–077–02　　　　**DOI：**10.19483/j.cnki.11−4653/n.2017.10.031

文 / 蒋文利

引言

大数据管理对广播媒体的意义不在于它的海量数据，而是在于对这些有用的信息进行媒体化的信息整理，也就是让大数据成为广播媒体的一种有价值的工具，通过对音频数据的深度挖潜，实现广播媒体资产的保值增值。广播媒体如何融入大数据时代，搭建大数据管理平台，利用对大数据的采集、内容的综合分析改变广播媒体的传播方式，创新营利模式，为广播节目的改版、名牌节目和名牌频率创新提供决策依据。探索出广播媒体在新的媒体环境下信息传播和媒体服务的发展之路，提升广播的核心竞争力，是广播人不可回避的问题。

1. 大数据技术特性、原理及关键技术

1. 1 大数据特点

大数据也称之为海量的资料，是利用现代化的 IT 技术产生的巨量信息资产，具有巨大的观察力、决策力、高成长性和信息种类多样性等特点。大数据管理是对不同的信息来源、不同格式的信息内容进行收集、存储和关联分析，从中获得新规律，创造出有新价值的信息技术和服务业态，它是数据容量巨大、种类繁多、索引和调取速度极快、可用价值高的巨型数据库。因此，大数据相比传统数据在“快速、价廉、优化”三个方面是最好的。

1. 2 大数据的关键技术

（1）大数据收集技术。是通过频道随录音频节目、网络广播及听众互动、社交网站互动所收集的结构化、非结构化和半结构化的巨量数据。利用两个功能模块实现大数据采集：大数据智能传感模块即数据传感部分、互联网数据交换部分、感知适配器部分、自动识别系统及软件硬件联接部分，对不同结构的巨量数据进行自动臻别、追踪、定位、上传下载、信号转码、监管、预处理等。大数据分析整理模块：将大数据平台的服务器虚拟化。采用分布式、网格节点数据存储方式，大数据提取、存储、集聚、解析和可视化接口技术，大数据的网络数据交换及数据压缩技术，大数据加密技术等完成对大数据的处理。

（2）大数据初始化技术。是对收集数据的辨别、分析、提取、整理等操作。重点包含两个方面。数据提取：可将这些不同结构类型的数据转化成统一的、方便处理的数据结构类型，提高数据解析和处理能力。数据清理：采集的大数据不是完全都有用的，首先要对采集的数据去伪存真和去粗取精，避免一些无用的信息对有用信息的干扰。数据清理的目的就是对有用和无用信息进行有效的过滤和分离，从而获得所需要的有用数据。

（3）大数据的存储和管理技术。大数据存储与监管就是把收集到的不同结构的数据保存到所建立的数据库中，并对其进行管理和上传下载。大数据存储方式有高速读存的存储、分布式文件存储、网格节点存储、去冗余数据存储及高性价比的大数据存储技术。大数据管理技术包括数据库建模、数据复合、不同类型大数据管理、大数据查寻及可视化等相关技术。

（4）大数据分析及挖掘技术。大数据分析提升了现有的音频节目数据高效利用，采取面向对象的数据联接、相似性联接大数据整合，解决网络媒体用户的主观好恶分析、在线收听分析等面向互联网的大数据分析技术。大数据挖潜技术是从庞大的、不完整的、不清晰的、随机的实际数据中，提取隐藏的、不可预知的，但又确实存在的有用信息内容的过程。

1. 3 构建大数据技术管理平台

广播的大数据管理需要建设大数据技术管理平台，来实现广播音频节目、文字稿件及相关的广播资源在平台上的联接、存储、解析、共建共享、索引调用和管理的目标。大数据平台包括音频节目联接模块、存储模块、解析模块和管理模块四个模块的内容。音频节目联接模块：它能完成最基本的技术服务功能，向电台各频道不同格式的整体节目和节目素材提供不同的类型的数据应用接口，为电台的音频节目大数据平台提供最基本的音频数据支撑。存储模块：它是大数据管理平台最主要的部分，是数据采集的归属地和资源池，用来存储不同格式的音频数据。采用音频的元数据和音频节目分开管理，查寻节目的元数据内容很少，占用的码流也就很少，实现音频节目高速查询功能。存储方式采用的是云存储结构，音频节目被存储在云存储系统中的不同存储节点上，音频节目是采用高速并行的访问方式，提高了节目的音频的上传和下载的速度。解析模块：它是大数据管理的核心模块，采用分布式存放和并行网格计算，并提供数据的聚集、类型划分、提供联接、测算和可视化插件及音频数据的深度挖潜、

解析、开发软件工具。管理模块：提供音频节目、元数据、编目、操作人员权限、数据接口、系统服务监控、平台运行维护等管理功能。

1.4 大数据技术原理

大数据管理不能用某一台电脑完成任务，应利用云计算的分布式、云端存储、虚拟化、高速并行数据库和互联网等相关技术，通过云计算系统平台向数十台、数百台乃至上千台计算机分配工作，并行计算和处理，然后再进行数据的汇总，最后得出对庞大数据进行分布式的挖掘后的运算结果。

大数据的技术原理包括以下几个方面：数据中心原理：对非结构音频数据处理从以“过程”为中心，转换成以“数据”为中心，通过多种技术对不同来源的信息进行整合，从而获得大数据。数据的价值原理：大数据的使用是通过互联网实现的，非网络环境 IT 产品使用体现在功能上；互联网环境 IT 产品使用体现在数据上。全标本原理：从过去的部分样本转变为全部样本，数据越是足够多，可视化程度越清晰、受众看得越清楚。注重效率原理：是由注重精准度变为注重效率，让广播媒体的发展决策更高效。注重关联性原理：从注重因果关系变为注重关联性，只需知道是什么，不用知道为什么。事件预测原理：大数据最重要价值是事件预测，是把云计算的网格运算运用到巨量的大数据上来，预知事件发生的可能性。

1.5 大数据的安全问题

之所以称之为“大数据”，是因为它的种类繁多、存储信息量庞大。它能把不同结构数据经过整理、提纯、解析后获得有利用价值的信息。大数据不仅广播媒体喜爱，黑客也在觊觎。大数据存在以下几方面的安全问题。云数据安全问题：大数据通常经过云端进行数据交互，极易遭到黑客的攻击，需要采用安全可靠的大数据云计算策略。网络防范：越来越多的线上交易、线上数据互换、线上互动都是在网络上进行的，使黑客得手的机会不断增加，一旦出现问题，大量的数据会丢失，媒体的声誉及经济利益遭受损失，甚至还会发生法律纠纷。大数据加密：大数据的采集、存储、数据在网络的上传下载，均通过计算机和移动设备，数据存在安全隐患，一旦信息泄露将导致不可挽回的损失，应利用透明加密技术对大数据加密。透明加密是使用者在透明加密环境下，系统采用智能方式对文件加密和解密，当退出透明加密环境时，黑客无法打开自动加密的文件，起到了对大数据的保护效果。

2. 大数据技术在广播中的应用

2.1 大数据为精准的收听率调查提供数据支撑

传统广播收听率调查是采用随机抽样的方式，对听众收听广播的地域、时间、收听工具等进行信息收集统计，它是一种概率统计结果，对数据内容的分析很少。利用大数据进行收听率的调查具有两个特点：一是不依照样本数据，采用的是总体数据。二是大数据在原理和应用上更加科学有效。所收集的数据大幅度增加，数据参数也不断更新，并对采集数据进行综合分析和深度挖掘，知其然更知所以然，不仅可以提高收听调查数据的精准性，所获得的数据具有丰富性、多样性和相关性，并整合出收听率调查数据库，降低获得收听率调查数据的成本，为电台的节目改版、节目创新提供强有力的数据支撑。

2.2 大数据使广播媒体获得的信息更加精准

互联网上聚集了数量庞大的各种信息，它包括音频节目、视频内容和文字稿件，而自媒体的内容更是取之不尽，每个网络用户都可以成为信息的发布者。大数据的采集渠道多样，数据种类繁多，数据量巨大，而这些数据是通过多个大型数据库进行数据解析、分类、归纳，编辑记者可从类型繁杂的数据中提取出对广播媒体有用的信息，使广播媒体获得的信息更加精准。

2.3 大数据助推广播媒体发挥作用及媒体资产增值

采用大数据管理对广播音频节目进行深层次的挖掘和管理，将使这些普通的音频节目产生具有数据功能的价值。电台媒体资源的真实性、公信力和权威性一直被社会所公认，对大数据的分析挖掘是在原始数据真实可信的基础上进行的。这些大数据可为广大受众和其他媒体提供新闻信息的增值服务。如：大数据可以对财经类新闻进行分类、整理、归纳，把整理出来的数据通过网络平台出售给客户和受众；让客户了解国家政策、经济走势、科技动向、市场行情，并进行市场预测和引导；扩大了网络自媒体的真实信息的来源，占领网络媒体的主阵地，让广播媒体发挥主流媒体应有的作用，避免虚假新闻在社会的负面影响。这样无疑可将电台的音频节目及相关数据得到多次应用，提升电台媒体资源的使用率和再利用率，让电台的媒体资产增值。

3. 结语

媒体在大数据时代的竞争就是媒体大数据的采集、整理、计算分析和应用能力强与弱的竞争。广播媒体只有掌握了大数据的采集、挖掘和分析，并从中获得对新闻内容更加深刻的观察、理解和预见，才能制作出一批大数据媒体节目。通过对大数据的管理和分析，对广播的发展格局做出相应的调整，为广播转型发展提供有价值的数据。广播媒体完全有能力成为大数据应用的引领者，为广播媒体未来取得长足发展探索出一条行之有效的发展之路。媒

参考文献

[1] 栾轶玫：《大数据重塑媒介生态》，《视听界》，2013（4）.

[2] 张守信 . 广电如何玩转大数据？——兼谈社交网络分析工具对广电大数据应用的启示 [J]. 声屏世界，2013（10）：8-11.

[3] 吴林飞 . 大数据时代的广电网络客户关系管理 [J]. 中国数字电视，2013（9）：39-42.

[4] 王建磊 . 大数据带给广电行业的本质变革 [J]. 声屏世界，2013（9）：8-11.

[5] 朱艳红 . 大数据时代广播电视节目的转型升级 [J]. 新闻研究导刊 .

（作者单位：天津广播电视台）